When Does $f^{-1} = 1/f$?

R. Cheng, A. Dasgupta, B. R. Ebanks, L. F. Kinch, L. M. Larson, and R. B. McFadden

1. INTRODUCTION. An unfortunate ambiguity in the standard notations for the reciprocal and the inverse of a function has led many students astray. It led us to ask when they actually do mean the same thing. To be more precise, suppose that f is a one-to-one function from some subset G of the real line onto itself. Then f has an inverse, and f^{-1} maps G onto itself. Assume further that G does not contain the origin, so that $1/f$ is defined throughout G. Can f have the property that

$$f^{-1} = 1/f? \tag{1}$$

Euler and Foran [3] have demonstrated that functions with property (1) do exist, but "are not ones likely to be encountered by a student in elementary calculus." They show, for instance, that a function on $(0, \infty)$ with property (1) must have an infinite number of discontinuities. Their results include piecewise continuous examples.

The present work concerns functions satisfying (1) whose domains are more general subsets of the real line $\mathbb{R}$, or of the complex plane $\mathbb{C}$. A complete description of the real solutions is given, along with examples showing how well- or poorly-behaved such functions can be. In the complex case, we consider analytic solutions to (1). It turns out that in contrast to the real case, there exist analytic solutions on simply connected regions in $\mathbb{C}$. We show that there is essentially only one meromorphic solution on the complex plane. Analytic local solutions are also characterized.

2. PRELIMINARIES. Throughout we shall consider functions on either $\mathbb{R}$ or $\mathbb{C}$. By $\mathrm{dom}(f)$, $\mathrm{ran}(f)$, and $\mathrm{gra}(f)$ we understand the domain, range, and graph, respectively, of the function f. If G is a subset of $\mathbb{R}$ or $\mathbb{C}$, then we define $-G = \{-x : x \in G\}$ and $1/G = \{1/x : x \in G\}$. At times we adopt the set-theoretic language in which a function is identified with its graph.

For a function f to have property (1), it is necessary to have $\mathrm{dom}(f) = \mathrm{ran}(f) = 1/\mathrm{dom}(f)$, and in particular, $\mathrm{dom}(f)$ cannot contain the origin. If (a, b) is a point in $\mathrm{gra}(f)$, then (b, a) is a point in $\mathrm{gra}(f^{-1})$. But property (1) requires that (b, a) belongs to $\mathrm{gra}(1/f)$, which implies that $(b, 1/a)$ belongs to $\mathrm{gra}(f)$. Continuing in this fashion, we see that a pair (a, b) is a member of $\mathrm{gra}(f)$ if and only if the points $(b, 1/a)$, $(1/a, 1/b)$, and $(1/b, a)$ are also members of $\mathrm{gra}(f)$. Thus, under iteration, a function f with property (1) acts in a cycle of order at most four:

$$\begin{tikzcd} a \arrow[r, "f"] & b \arrow[d, "f"] \\ \frac{1}{b} \arrow[u, "f"] & \frac{1}{a} \arrow[l, "f"] \end{tikzcd} \tag{2}$$

On the other hand, if a set of ordered pairs can be partitioned into subsets of the form $\{(a, b), (b, 1/a), (1/a, 1/b), (1/b, a)\}$, then it is the graph of a solution to (1), provided only that it is the graph of a function. This phenomenon was already observed in [**3**].

The cyclic behavior can be understood in another way. Indeed, if f satisfies (1) then

$$f(1/f(x)) = x$$

and

$$1/f(f(x)) = x \tag{3}$$

for all x; these are both equivalent to condition (1). Since (3) can be written

$$f(f(x)) = 1/x, \tag{4}$$

it is clear that $f \circ f \circ f \circ f$ is the identity function. This perspective gives condition (1) a grounding in the functional equations literature; particularly that devoted to iterative roots of the identity such as [**4**, **5**].

3. REAL SOLUTIONS. We first consider the case $\mathrm{dom}(f) \subseteq \mathbb{R}$.

Here is a way to construct all solutions of (1) on $(0, \infty)$ (or with appropriate modifications, any other suitable subset of $\mathbb{R}$). It is adapted from [**4**], which contains more general results. The idea, also mentioned in [**3**], is to impose the 4-cycle behavior (2) upon f, after having selected the points to serve in the roles of a and b.

First, it must be that $f(1) = 1$. For if $f(1) = b$, then $f(b) = 1/1 = 1$, and hence $f(1) = 1/b$; now $b = 1/b$ in $(0, \infty)$ forces $b = 1$. Next, partition $(0, \infty) \setminus \{1\}$ into equivalence classes under the relation $x \sim x$, $x \sim 1/x$. Partition the resulting space S into two subsets U and V of the same cardinality. Let $g: U \mapsto V$ be an arbitrary bijection, and define $\phi : S \mapsto S$ by

$$\phi(s) = \begin{cases} g(s), & \text{if } s \in U \\ g^{-1}(s), & \text{if } x \in V \end{cases} \tag{5}$$

Observe that $\phi \circ \phi$ is the identity on $S = U \cup V$.

Let C be an arbitrary subset of $(0, \infty) \setminus \{1\}$ that contains exactly one member of each equivalence class $\{x, 1/x\}$ in S; the Axiom of Choice is needed for this step. For each $x \in (0, \infty) \setminus \{1\}$, let $[x]$ denote the representative of $\{x, 1/x\}$ that belongs to C. Define f from $(0, \infty)$ to $(0, \infty)$ by

$$f(x) = \begin{cases} [\phi(\{x, 1/x\})], & \begin{array}{l}\text{if } \{x, 1/x\} \in U \text{ and } [x] = x, \\ \text{or if } \{x, 1/x\} \in V \text{ and } [x] = 1/x\end{array} \\ 1, & \text{if } x = 1 \\ 1/[\phi(\{x, 1/x\})], & \begin{array}{l}\text{if } \{x, 1/x\} \in V \text{ and } [x] = x, \\ \text{or if } \{x, 1/x\} \in U \text{ and } [x] = 1/x.\end{array} \end{cases} \tag{6}$$

It is straightforward to check that this function satisfies (1).

Theorem 3.1. *Every solution of* (1) *on* $(0, \infty)$ *is obtained by the preceding construction.*

Proof: Suppose that f satisfies (1) on $(0,\infty)$. Partition $(0,\infty)\setminus\{1\}$ using the 4-cycles $\{x, f(x), 1/x, 1/f(x)\}$ of f; a cycle of order 2 or 1 can arise only from $f(1) = 1$. Let E contain exactly one member from each 4-cycle. Define $U = \{\{x, 1/x\} : x \in E\}$ and $V = \{\{f(x), 1/f(x)\} : x \in E\}$. Then $S = U \cup V$ is the set of all pairs $\{x, 1/x\}$ from $(0,\infty)\setminus\{1\}$. Note also that f induces an invertible mapping g from U onto V by $g(\{x, 1/x\}) = \{f(x), 1/f(x)\}$. Take $C = E \cup f(E)$, noting that C selects precisely one member of each pair $\{x, 1/x\}$, $x \in (0,\infty)\setminus\{1\}$. With U, V, S, g, and C thus defined, and with ϕ defined as in (5), the function in (6) is indeed the original function f. To check this, let $a \in E$ and $f(a) = b$. Then $\{a, 1/a\} \in U, \{b, 1/b\} \in V$, and $\{a, b\} \subseteq C$. Thus, the constructed function (6) produces the 4-cycle (2), in agreement with the given function f. ■

It is clear that most solutions of (1) constructed in this way are uninteresting. The solutions become interesting only if they interact with some other feature (e.g., the topology) of the domain.

4. PIECEWISE CONTINUOUS SOLUTIONS. We turn to the question of continuity. If (1) holds, it cannot happen that f is monotone increasing throughout its domain; for then f^{-1} must be monotone increasing and $1/f$ monotone decreasing. In a similar way we see that f cannot be monotone decreasing. Thus (1) forces $\mathrm{dom}(f)$ to be disconnected, or f to be discontinuous.

Let us examine the functions satisfying condition (1) whose restrictions to certain intervals are continuous. We call a finite sequence $(I_1, I_2, \ldots, I_n)$ of pairwise disjoint nondegenerate intervals an *n-circuit* of f, if $\bigcup_{k=1}^{n} I_k \subseteq \mathrm{dom}(f)$, f restricted to each I_k is continuous, $f(I_k) = I_{k+1}$ for all $k \in \{1, 2, \ldots, n-1\}$, and $f(I_n) = I_1$.

Lemma 4.1. *Suppose f satisfies* (1). *If $n \notin \{2, 4\}$, then f does not have an n-circuit.*

Proof: Suppose (I_1) is a 1-circuit for f. Then the restriction of f to I_1 is continuous and satisfies (1). But this possibility has already been ruled out.

Next, assume (I_1, I_2, I_3) is a 3-circuit. Then $I_3 = f^{-1}(I_1) = 1/f(I_1) = 1/I_2$, and $I_2 = f^{-1}(I_3) = 1/f(I_3) = 1/I_1$. This gives $I_1 = 1/I_2 = I_3$, contradicting disjointness.

Finally, if there is an n-circuit $(I_1, I_2, \ldots, I_n)$ with $n \geq 4$, then we have $I_n = f^{-1}(I_1) = 1/f(I_1) = 1/I_2$. On the other hand, $I_2 = f^{-1}(I_3) = 1/f(I_3) = 1/I_4$. It follows that $I_n = 1/I_2 = I_4$, forcing $n = 4$. ■

Let us say that a function f is *piecewise continuous* if $\mathrm{dom}(f)$ is the disjoint union of intervals $\{J_1, J_2, J_3, \ldots\}$ such that the restriction of f to each J_k is continuous. We insist that each J_k be maximal in the sense that if $J_k \subseteq J$ for some interval J, and f is continuous on J, then $J_k = J$. Assume for simplicity that each J_k is also nondegenerate.

Piecewise continuous solutions of (1) enjoy the following characterization.

Theorem 4.2. *Suppose f is a piecewise continuous function, and let $\mathcal{J} = \{J_1, J_2, J_3, \ldots\}$ be the associated sequence of maximal intervals of continuity. Then f satisfies* (1) *if and only if $\mathcal{J}$ can be partitioned into $\{\mathcal{J}_k : k = 1, 2, \ldots\}$ such that for each k, $\mathcal{J}_k$ is a 2- or 4-circuit for f, $f|_{\cup \mathcal{J}_k}$ is continuous, and $f|_{\cup \mathcal{J}_k}$ satisfies* (1).

Proof: It is easy to see that if f has the stated structure, then f satisfies (1).

On the other hand, suppose that f satisfies (1). Then for each k, the connected set J_k is mapped to an interval. Now f^{-1} is also piecewise continuous, with the same sequence of maximal intervals. It follows that $f(J_k)$ must be a member of $\mathscr{J}$. By Lemma 4.1, either $(J_k, f(J_k))$ is a 2-circuit for f, or $(J_k, f(J_k), [f \circ f](J_k), [f \circ f \circ f](J_k))$ is a 4-circuit for f. ■

Simply put, a piecewise continuous solution is a union of continuous solutions on disjoint n-circuits, where n is 2 or 4. It remains to describe those continuous functions with property (1) defined over a single n-circuit.

A continuous solution f on a 2-circuit can be constructed on intervals $I_1 \subseteq (0, \infty)$ and $I_2 \subseteq (-\infty, 0)$, such that $I_1 = 1/I_1$ and $I_2 = 1/I_2$. Choose $f_1 : I_1 \mapsto I_2$ to be continuous and monotone, with $f_1(1) = -1$, but otherwise arbitrary. Define $f_2 : I_2 \mapsto I_1$ by $f_2(x) = 1/f_1^{-1}(x)$, $x \in I_2$. Then $f = f_1 \cup f_2$ satisfies (1) on the 2-circuit (I_1, I_2).

On the other hand, the necessity of $I_1 = 1/I_1$ follows from $I_1 = f(I_2) = f(f(I_1)) = 1/I_1$. This, in turn, forces $\operatorname{dom}(f)$ to contain both 1 and -1, whereupon we may choose $1 \in I_1 \subseteq (0, \infty)$ and $-1 \in I_2 \subseteq (-\infty, 0)$.

A continuous solution f to (1) on a 4-circuit can be built upon $(A, B, 1/A, 1/B)$, where A and B are disjoint intervals, neither containing -1, 0, or 1. Choose $f_1 : A \mapsto B$ to be continuous and monotone, but otherwise arbitrary. Define $f_2 : B \mapsto 1/A$ by $f_2(x) = 1/f_1^{-1}(x)$, $x \in B$; define $f_3 : 1/A \mapsto 1/B$ by $f_3(x) = 1/f_2^{-1}(x)$, $x \in 1/A$; and finally define $f_4 : 1/B \mapsto A$ by $f_4(x) = 1/f_3^{-1}(x)$, $x \in 1/B$. Then $f = f_1 \cup f_2 \cup f_3 \cup f_4$ is continuous and satisfies (1) on $A \cup B \cup 1/A \cup 1/B$. Every continuous solution on a 4-circuit arises in this way.

Note that every 2-circuit gives rise to a 4-circuit when the points $(1, -1)$ and $(-1, 1)$ are deleted from the graph of f.

Example 4.3. *A continuous solution to* (1) *on* $\mathbb{R} \setminus \{0\}$ *with a* 2-*circuit*:

$$f(x) = \begin{cases} -x, & \text{if } x \in (0, \infty) \\ -1/x, & \text{if } x \in (-\infty, 0). \end{cases}$$

Example 4.4. *A continuous solution of* (1) *with a* 4-*circuit*; *the domain of f is disconnected*, *as is necessary for continuity*:

$$f(x) = \begin{cases} 2x, & \text{if } x \in [2, 3] \\ 2/x, & \text{if } x \in [4, 6] \\ x/2, & \text{if } x \in [1/3, 1/2] \\ 1/2x, & \text{if } x \in [1/6, 1/4]. \end{cases}$$

Let $\mathbb{N}$ be the set of natural numbers.

Example 4.5. *A solution to* (1) *on* $(0, \infty)$:

$$f(x) = \begin{cases} 1/(x+1), & \text{if } x \in (2n-1, 2n],\ n \in \mathbb{N} \\ x - 1, & \text{if } x \in (2n, 2n+1],\ n \in \mathbb{N} \\ 1, & \text{if } x = 1 \\ (x+1)/x, & \text{if } x \in [1/2n, 1/(2n-1)),\ n \in \mathbb{N} \\ x/(1-x), & \text{if } x \in [1/(2n+1), 1/2n),\ n \in \mathbb{N}. \end{cases}$$

The domain of f is the disjoint union of $\{1\}$ *and a sequence of* 4-*circuits*. *The domain is connected*, *and f has infinitely many discontinuities*. This is [**3**].

Example 4.6. *A solution to* (1) *with only* 3 *discontinuities in* $(0, \infty)$:

$$f(x) = \begin{cases} 1/(x + \frac{1}{2}), & \text{if } x \in (0, \frac{1}{2}) \\ x - \frac{1}{2}, & \text{if } x \in (\frac{1}{2}, 1) \\ 1, & \text{if } x = 1 \\ 2x/(2 - x), & \text{if } x \in (1, 2) \\ (x + 2)/2x, & \text{if } x \in (2, \infty). \end{cases}$$

This is possible since $1/2$ and 2 are excluded from the domain. A similar effect is illustrated in [**3**, Example 3], where the points 0 and ∞ are adjoined to the domain.

5. AN EXAMPLE WITH DENSE GRAPH. The examples of Section 4 give us a sense of how well-behaved the solutions of (1) can be. Now we present an example that illustrates the opposite extreme.

Proposition 5.1. *There exists a function f that satisfies* (1) *on* $(0, \infty)$ *such that* $\text{gra}(f)$ *is dense in* $(0, \infty) \times (0, \infty)$.

Proof: We shall construct the graph of such a function f. The strategy is to choose a dense set of points to serve the role of a in the 4-cycle (2), then impose the 4-cycle to get property (1).

Let $\{E_k\}$ be a relatively open countable basis of $(0,1) \times (0,1)$. Let $(a_1, b_1) \in E_1$ be chosen arbitrarily. Let $F_1 = \{(a_1, b_1), (b_1, 1/a_1), (1/a_1, 1/b_1), (1/b_1, a_1)\}$. Having defined F_{n-1}, choose $(a_n, b_n) \in E_n$ so that neither a_n nor b_n lies in $\bigcup_{k=1}^{n-1} \{a_k, b_k, 1/a_k, 1/b_k\}$. Let $F_n = F_n \cup \{(a_n, b_n), (b_n, 1/a_n), (1/a_n, 1/b_n), (1/b_n, a_n)\}$. Then the function $f_1 = \bigcup_{k=1}^{\infty} F_k$ satisfies (1) on a countable subset of $(0, \infty)$, and its graph is dense in the first quadrant.

There are uncountably many points remaining in $(0, \infty) \setminus \text{dom}(f_1)$. Define f_2 to satisfy (1) on $(0, \infty) \setminus \text{dom}(f_1)$ using the scheme from Section 3. Then $f = f_1 \cup f_2$ satisfies (1) on $(0, \infty)$, and its graph is dense in the first quadrant. ∎

With minor changes, this construction can be used to produce an example of a function on $\mathbb{R} \setminus \{0\}$ that satisfies (1) and whose graph is dense in $\mathbb{R} \times \mathbb{R}$.

6. COMPLEX SOLUTIONS. Let us next investigate analytic solutions of (1) on open subsets of the complex plane. Many of the examples from the real case extend in a simple way to this setting.

Example 6.1. *Let* $G_1 = \{z \in \mathbb{C} : \operatorname{Re} z > 0\}$ *and* $G_2 = \{z \in \mathbb{C} : \operatorname{Re} z < 0\}$. *Define f on* $G = G_1 \cup G_2$ *by*

$$f(z) = \begin{cases} -1/z, & \text{if } z \in G_1 \\ -z, & \text{if } z \in G_2. \end{cases}$$

Then (1) *holds and f is analytic on G. This function extends Example* 4.3.

Example 6.2. *Let* G_1, G_2, G_3, *and* G_4 *be the open circular regions in* $\mathbb{C}$ *such that the real intervals* $(1, 2)$, $(2, 3)$, $(1/2, 1)$, *and* $(1/3, 1/2)$ *are respective diametric chords.*

Define f on $G = G_1 \cup G_2 \cup G_3 \cup G_4$ *by*

$$f(z) = \begin{cases} 1/(z+1), & \text{if } z \in G_1 \\ z-1, & \text{if } z \in G_2 \\ (z+1)/z, & \text{if } z \in G_3 \\ z/(1-z), & \text{if } z \in G_4. \end{cases}$$

Then f satisfies (1) *and is analytic on G. This function corresponds to Example* 4.5, *restricted to a* 4*-circuit of open sets.*

Are there analytic solutions that are fundamentally different from real solutions? We begin by exploring how large the domain of a complex solution can be. First, a function satisfying (1) cannot be entire, as its domain must exclude the origin. But much more can be said.

Theorem 6.3. *Let G be the complex plane, excluding a collection of isolated points. If f is analytic and satisfies* (1) *on G, then $G \subseteq \mathbb{C} \setminus \{-i, 0, i\}$, and f or f^{-1} is given by*

$$z \mapsto \frac{z+i}{iz+1}. \tag{7}$$

This asserts that, perhaps surprisingly, there is really only one pair of solutions f and f^{-1} to equation (1) that are meromorphic on $\mathbb{C}$; other such solutions are restrictions of these two. Note that their domain is connected; in contrast, real solutions on intervals must have discontinuities.

The mechanism underlying this example can be understood geometrically. Under the mapping (7), the points in the right halfplane are given a sort of non-Euclidean quarter turn clockwise about the fixed point 1. In a symmetric way, the points of the left halfplane are given a quarter turn counterclockwise about -1. Points belonging to the imaginary axis are shifted upward, so that the singularities are permuted cyclically:

$$0 \mapsto i \mapsto \infty \mapsto -i \mapsto 0.$$

It follows that $z \mapsto f(f(z))$ gives rise to half turns about the points 1 and -1, and the singularities are mapped as follows:

$$0 \leftrightarrow \infty \quad \text{and} \quad i \leftrightarrow -i.$$

But that is also what the mapping $z \mapsto 1/z$ does, and so condition (1) is implemented in the form $f(f(z)) = 1/z$.

The proof of Theorem 6.3 relies on a chain of lemmas. Let us say that the point a is an *antipole* of f if a is a zero of f, or if a is a removable singularity of f such that $\lim_{z \to a} f(z) = 0$.

Lemma 6.4. *Let G be the complex plane, excluding a collection of isolated points. If f is analytic on G and satisfies* (1), *then f is not the restriction to G of an entire function.*

Proof: By [**2**, XII.4.4], there are two mutually exclusive possibilities for an entire function: it is a polynomial, or it attains uncountably many values infinitely often. In the latter case, exclusion of a collection of isolated points from the domain cannot produce a univalent function. Moreover, most polynomials can be ruled out for having more than one distinct root. Otherwise, consider f of the form

$$f(z) = a(z-w)^m,$$

where $a \in \mathbb{C}$, $w \in \mathbb{C}$ and m is a nonnegative integer. One verifies by inspection that f cannot satisfy $f(1/f(z)) = z$. ∎

Lemma 6.5. *Suppose that a nonconstant analytic function f has removable singularities a_1 and a_2 such that $\lim_{z\to a_1} f(z) = w_0 = \lim_{z\to a_2} f(z)$. Then for all w in a neighborhood of w_0, f attains the value w at least twice.*

Proof: Let g be the analytic extension of f to $G \cup \{a_1, a_2\}$. Let U_1 and U_2 be disjoint open balls in G, centered at a_1 and a_2, respectively. Then $V = g(U_1) \cap g(U_2)$ is an open neighborhood of w_0. It is clear that for every $w \in V \setminus \{w_0\}$, $f^{-1}(w)$ contains at least two distinct points. ■

Corollary 6.6. *Let G be the complex plane, excluding a collection of isolated points. If f is analytic on G and satisfies* (1), *then f has at most one antipole.*

Lemma 6.7. *Let G be the complex plane, excluding a collection of isolated points. If f is analytic on G and satisfies* (1), *then f cannot have an essential singularity.*

Proof: The Great Picard Theorem [**2**, XII.4.2] implies that in every neighborhood of an essential singularity, the function attains an uncountable number of values infinitely often. No restriction of such a function by excluding isolated points can satisfy (1). ■

Lemma 6.8. *Let G be the complex plane, excluding a collection of isolated points. If f is analytic on G and satisfies* (1), *then f has precisely one pole and one antipole in the complex plane.*

Proof: If f satisfies (1) and has no antipole, then $1/f$ satisfies (1) and has no poles or essential singularities. Thus $1/f$ extends to an entire function, in violation of Lemma 6.4. It follows that f must have at least one antipole, and by Corollary 6.6, at most one as well. Repeating this argument with $f^{-1} = 1/f$, we see that f has exactly one pole. ■

To complete the proof of Theorem 6.3, let G be the plane excluding a collection of isolated points. Then f satisfying (1) on G must be of the form

$$f(z) = \frac{z-a}{z-b} \cdot \phi(z),$$

where ϕ is entire and nonvanishing. From $f^{-1} = 1/f$ we get

$$f^{-1}(z) = \frac{z-b}{z-a} \cdot \frac{1}{\phi(z)}.$$

Thus a is an antipole of f, and b is an antipole of f^{-1}. This in turn gives

$$\lim_{z\to 0} f(z) = b = \frac{-a}{-b} \cdot \phi(0)$$

and

$$\lim_{z\to 0} f^{-1}(z) = a = \frac{-b}{-a} \cdot \frac{1}{\phi(0)}.$$

It follows that $\phi(0) = a^{-3} = b^3$.

If $\lim_{z\to 0} \phi(1/z)$ exists, then ϕ is bounded and entire, hence constant, according to Liouville's theorem. Otherwise, the point at infinity is a pole or an essential singularity of ϕ. In the former case, $1/\phi$ is bounded and entire, and again ϕ is

constant. The latter case cannot occur, for then f would also have an essential singularity at infinity, and hence could not be one-to-one. Thus we find that ϕ must be constant. In conclusion, f must be of the form

$$f(z) = \frac{1}{a^2} \cdot \frac{z-a}{az-1}.$$

By insisting that f satisfy (1) we get $a = \pm i$. This verifies the claim in Lemma 6.4.

7. SOLUTIONS ON SIMPLY CONNECTED REGIONS. By a *region* in $\mathbb{C}$ we mean a nonempty, open, connected subset. A region G is *simply connected* if every closed curve in G can be continuously deformed into a point; informally, G has no "holes." We shall exploit simple connectedness through several technical tools from complex analysis, including the following, known as the Riemann Mapping Theorem.

Theorem 7.1. *Suppose G is a simply connected region that is not the whole complex plane, and let $a \in G$. There exists a unique function ϕ such that ϕ is analytic, ϕ is a bijection of the unit disc D onto G, $\phi(0) = a$, and $\phi'(0) > 0$.*

For a proof, see [**2**, VII.4.2]. The function ϕ is called a Riemann mapping function for G. It allows us to reduce some questions concerning G to the familiar disc D.

Using this approach, we find that there are analytic solutions on simply connected domains, all arising from a mechanism that is peculiar to the geometry of $\mathbb{C}$.

Theorem 7.2. *Suppose G is a simply connected region. There exists an analytic function f satisfying* (1) *on G if and only if $G = 1/G$. In this case, G must contain exactly one of the points* 1 *or* -1.

If G contains 1, *let* $L(z) = \log z$, *where* log *is the branch of the logarithm on G such that* $\log 1 = 0$; *if G contains* -1, *let* $L(z) = \log(-z)$, *where* log *is the branch of the logarithm on* $-G$ *such that* $\log 1 = 0$. *Then*

$$f = L^{-1} \circ \phi \circ (\pm i \cdot \phi^{-1}) \circ L, \tag{8}$$

where ϕ is the unique Riemann mapping function for $L(G)$ satisfying $\phi(0) = 0$.

Proof: Let G be a simply connected region that is not the whole complex plane. For f to satisfy (1) on G, the domains of f and $1/f$ must coincide, hence it is necessary that $G = 1/G$.

Suppose $G = 1/G$. Since G is simply connected, there is a branch $\lambda(z)$ of the logarithm on G. Let $V = \lambda(G)$. Then V is simply connected; indeed, suppose σ is any smooth closed curve whose trace lies in V, and let F be any analytic function on V. Then,

$$\oint_{\sigma} F(z)\,dz = \oint_{\lambda^{-1}\circ\sigma} F(\lambda(w))\,\lambda'(w)\,dw = 0$$

because G is simply connected. Furthermore, $G = 1/G$ implies that $V = -V + 2\pi ik$ for some integer k. We claim that V contains πik. To see this, let a be any point in V. If $a = \pi ik$, then there is nothing more to show. Otherwise, V also contains $-a + 2\pi ik$. By simple connectedness, there is a simple path γ connecting a to $-a + 2\pi ik$. By its symmetry property, V must also contain the

path $\delta = -2\gamma + \pi ik$, which connects $-a + \pi ik$ to a. If either γ or δ should contain πik itself, then there is nothing more to show. Otherwise, let Λ be a branch of the logarithm defined in a neighborhood of the path γ, and observe that

$$\oint_{\gamma\cup\delta} \frac{1}{z - \pi ik}\, dz = \int_{\gamma} \frac{1}{z - \pi ik}\, dz - \int_{\gamma} \frac{1}{-z + \pi ik}\, dz$$
$$= 2[\Lambda(a - \pi ik) - \Lambda(-a + \pi ik)] = \pm 2\pi i,$$

depending on the orientation of $\gamma \cup \delta$. This shows that $\gamma \cup \delta$ winds around the point πik exactly once. Since V is simply connected, it must contain the point πik after all, and hence G contains 1 or -1. A similar argument shows that if G were to contain both 1 and -1, then it would also contain 0, which cannot happen.

Now assume that there exists a function f on G satisfying (1). Let $L(z)$ be defined as in the statement of the theorem. Let U be the open set $L(G)$, and define $\psi : L(G) \mapsto L(G)$ by

$$\psi = L \circ f \circ L^{-1}. \tag{9}$$

Then ψ is univalent and bijective on U, and $\psi(0) = 0$. Furthermore, (4) ensures that

$$\begin{aligned}\psi(\psi(z)) &= [L \circ f \circ L^{-1}] \circ [L \circ f \circ L^{-1}](z)\\ &= [L \circ f \circ f \circ L^{-1}](z) = L(1/L^{-1}(z))\\ &= -L(L^{-1}(z)) = -z.\end{aligned}$$

Since $\psi'(\psi(z))\psi'(z) = -1$, taking $z = 0$ gives $\psi'(0)^2 = -1$.

Note that U is a simply connected region. Hence there is a unique Riemann mapping function ϕ from D onto U such that $\phi(0) = 0$. Thus defined, ϕ is an odd function. Indeed, since $G = 1/G$, we have $U = L(G) = L(1/G) = -L(G) = -U$; and now both $-\phi(-z)$ and $\phi(z)$ are Riemann mapping functions for U that map 0 to itself. By uniqueness, they must coincide.

Define $\tau = \phi^{-1} \circ \psi \circ \phi$. Note that τ is a bijective analytic function on D, $\tau(0) = 0$, and

$$\begin{aligned}\tau'(0) &= [(\phi^{-1})' \circ \psi \circ \phi](0) \cdot [\psi' \circ \phi](0) \cdot \phi'(0)\\ &= (1/\phi'(0)) \cdot \psi'(0) \cdot \phi'(0) = \psi'(0).\end{aligned}$$

That is, $\overline{\psi'(0)}\tau$ is a Riemann mapping function for D that preserves the origin. But the identity function is also a Riemann mapping function on D preserving the origin. It must be that $\tau(z) = \psi'(0)z$ $(= \pm iz)$. And now by (9),

$$\begin{aligned}f &= L^{-1} \circ \psi \circ L\\ &= L^{-1} \circ \phi \circ \phi^{-1} \circ \psi \circ \phi \circ \phi^{-1} \circ L\\ &= L^{-1} \circ \phi \circ \tau \circ \phi^{-1} \circ L.\end{aligned}$$

Thus f has the asserted structure.

On the other hand, suppose $G = 1/G$ and define f by equation (8). Note that

$$\begin{aligned}f(G) &= [L^{-1} \circ \phi \circ \tau \circ \phi^{-1} \circ L](G)\\ &= [L^{-1} \circ \phi \circ \tau \circ \phi^{-1}](U)\\ &= [L^{-1} \circ \phi \circ \tau](D) = [L^{-1} \circ \phi](iD)\\ &= [L^{-1} \circ \phi](D) = L^{-1}(U) = G,\end{aligned}$$

where $\tau(z) = \pm iz$, so f does indeed map G onto itself. Finally, check that

$$\begin{aligned} f(f(z)) &= \left[L^{-1} \circ \phi \circ \tau \circ \phi^{-1} \circ L \circ L^{-1} \circ \phi \circ \tau \circ \phi^{-1} \circ L\right](z) \\ &= \left[L^{-1} \circ \phi \circ \tau \circ \tau \circ \phi^{-1} \circ L\right](z) \\ &= L^{-1}\left(\phi\left(-1 \cdot \phi^{-1}(L(z))\right)\right) = L^{-1}(-L(z)) \\ &= L^{-1}(L(1/z)) = 1/z, \end{aligned}$$

proving that (1) holds. ■

The choice of sign in (8) merely distinguishes f from f^{-1}. Thus a pair f and f^{-1} of analytic solutions to (1) are uniquely determined by a simply connected region G satisfying $G = 1/G$.

An immediate example arises from taking ϕ to be the identity on D.

Example 7.3. *Let $G = \exp D$, and define $f(z) = \exp(i \log z)$, where* log *is the principal branch of the logarithm, restricted to G. Then f satisfies* (1) *on G.*

This example illustrates the mechanism underlying (1) when G is simply connected. Here, $\log \circ f \circ \exp$ is just multiplication by i on the disc D, that is, a quarter turn about the origin. Thus, f itself gives rise to a non-Euclidean quarter turn about the point 1. It follows that $z \mapsto f(f(z))$ is something like a half turn around the point 1, in agreement with the mapping $z \mapsto 1/z$. As with the meromorphic example (7), a quarter turn mechanism here brings about condition (1) in the form $f(f(z)) = 1/z$. The general construction from Theorem 7.2 merely extends this effect from D to $\log G$ by factoring through a Riemann mapping function.

If G is the right half plane, the function constructed in (8) coincides with the meromorphic solution (7), restricted to G.

The image of the right half plane under the invertible mapping $z \mapsto z^2$ is a slit plane, the "largest" possible simply connected region admitting a solution to (1).

Example 7.4. *Let $G = \mathbb{C} \setminus \{z : z \le 0\}$, and define*

$$f(z) = \left[\frac{z^{1/2} + i}{iz^{1/2} + 1}\right]^2,$$

where $z^{1/2}$ is the principal branch of the square root on G. Then f satisfies (1) *on G. Here $\phi(z) = 2\log[(1 + z)/(1 - z)]$ is the associated Riemann mapping function.*

For multiply connected regions, the issue is more complicated. It is clear that some examples can be constructed from those on simply connected regions by excluding certain points from the domain. For example, let $K = \{z : |z| \le 1/2\}$, and consider the function f of Example 7.3 restricted to $G_1 = \exp(D \setminus K)$. Then the restricted function is analytic and satisfies (1) on the region G_1, which is homeomorphic to an annulus.

But the following example shows that there is an annulus G on which there is no analytic solution of (1), even though $G = 1/G$.

Proposition 7.5. *Let G be the annulus $\{z : 1/2 < |z| < 2\}$. There is no analytic function on G satisfying* (1).

Proof: Suppose there is an analytic solution f of (1) on G. Informally, we see that either f preserves the two components of the boundary of G (in some limiting sense), or it reverses them. Either way, the composite $f \circ f$ must then preserve them, contrary to what the mapping $z \mapsto 1/z$ does.

To carry out this argument, define $C_r = \{z : |z| = r\}$ for $1/2 < r < 2$, and consider $f(C_r)$. Note that $1/C_r = C_{1/r}$, and $f^{-1}(C_r) = 1/f(C_r)$. By examining the winding number

$$\frac{1}{2\pi i}\oint_{f(C_r)} \frac{1}{z}\, dz = \frac{1}{2\pi i}\oint_{C_r} \frac{(f^{-1})'(w)}{f^{-1}(w)}\, dw,$$

we see that each $f(C_r)$ is a smooth closed curve winding once around the hole. Since $(f^{-1})'/f^{-1}$ is analytic in G, the second integral would be zero if C_r could be deformed to a point. Thus it must wind around the hole some nonzero number of times. But then, by univalence, the path $f(C_r)$ does not intersect itself (except at the endpoints), so the winding number must be exactly 1 or -1.

It follows that $f(C_r)$ can be continuously deformed to either C_r or $-C_r$ inside G, respectively; hence, $(f \circ f)(C_r)$ must have winding number 1 about the origin. On the other hand, from property (1), we have

$$\begin{aligned} 1 &= \frac{1}{2\pi i}\oint_{(f \circ f)(C_r)} \frac{1}{z}\, dz = \frac{1}{2\pi i}\oint_{1/C_r} \frac{1}{z}\, dz \\ &= -\frac{1}{2\pi i}\oint_{C_{1/r}} \frac{1}{z}\, dz = -1, \end{aligned}$$

a contradiction. Thus, no such function f exists. ∎

8. ANALYTIC LOCAL SOLUTIONS. Let ϕ be a given function. A solution f of the functional equation $f \circ f = \phi$ is said to be an *iterative square root* of ϕ. A function satisfying (1) has the equivalent property

$$f(f(z)) = 1/z, \tag{10}$$

which is to say that $f(z)$ is an iterative square root of the function $1/z$. A general treatment of analytic local iterative roots of the identity can be found in **[5]**. We now apply similar ideas to equation (10).

Theorem 8.1. *Let g be analytic in a neighborhood of* 1, *and suppose* $g'(1) \neq 0$. *Let* $a^2 = -1$, *and define*

$$\sigma(z) = g(z) - g(1/z). \tag{11}$$

The function

$$f(z) = \sigma^{-1}(a \cdot \sigma(z)) \tag{12}$$

is an analytic local iterative square root of $1/z$ *with* $f(1) = 1$.

Conversely, if $f(z)$ is an analytic local iterative square root of $1/z$ *and* $f(1) = 1$, *then f has the representation* (12).

Proof: Let g be analytic in a neighborhood of 1, and suppose $g'(1) \neq 0$. Let $a^2 = -1$, and define $\sigma(z) = g(z) - g(1/z)$. Then

$$\sigma'(z) = g'(z) + (1/z^2)g'(1/z),$$

so $\sigma'(1) = 2g'(1)$ is nonzero. Hence σ^{-1} exists in a neighborhood of $\sigma(1) = g(1) - g(1) = 0$. With f defined as in (12), and observing that

$$\sigma(1/z) = g(1/z) - g(z) = -\sigma(z),$$

we have $\sigma(1) = 0$, $f(1) = \sigma^{-1}(a \cdot \sigma(1)) = \sigma^{-1}(0) = 1$, and

$$\begin{aligned} f(f(z)) &= \sigma^{-1} \circ (a\sigma) \circ \sigma^{-1} \circ (a\sigma)(z) \\ &= \sigma^{-1} \circ (a^2\sigma)(z) = \sigma^{-1}(-\sigma(z)) = 1/z. \end{aligned}$$

Thus f is an analytic solution to (10) near 1.

Conversely, suppose f is an analytic local iterative square root of $1/z$, and $f(1) = 1$. Put $a = f'(1)$. Then $f'(f(z)) \cdot f(z) = (d/dz)(1/z) = -1/z^2$, so $a^2 = f'(1)^2 = f'(f(1)) \cdot f'(1) = -1$ as claimed. Now

$$g(z) = z + a^{-1}f(z)$$

defines an analytic function g in a neighborhood of 1, with

$$g'(1) = 1 + a^{-1}f'(1) = 2,$$

which in particular is nonzero. Next, define σ as in (11), so that

$$\sigma(z) = z + a^{-1}f(z) - z^{-1} - a^{-1}f(1/z),$$

and $\sigma'(1) = 4$ (again, nonzero). Consequently,

$$\begin{aligned} \sigma(f(z)) &= f(z) + a^{-1}f(f(z)) - (1/f(z)) - a^{-1}f(1/f(z)) \\ &= f(z) + a^{-1}z^{-1} - f(1/z) - a^{-1}z \\ &= a\left[a^{-1}f(z) + a^{-2}z^{-1} - a^{-1}f(1/z) - a^{-2}z\right] \\ &= a\left[a^{-1}f(z) - z^{-1} - a^{-1}f(1/z) + z\right] = a\sigma(z). \end{aligned}$$

Therefore for z sufficiently close to 1, the function $f(z)$ is of the form prescribed in (12). ■

This approach has the advantage of being elementary, at the expense of generally yielding only local information. The mechanism underlying these local solutions is very similar to the quarter turn effect, which we have already encountered. To see this, suppose that f is an analytic local solution with $f(1) = 1$, and let f have the power series

$$f(z) = 1 + c_1(z-1) + c_2(z-1)^2 + c_3(z-1)^3 + \cdots$$

for z near 1. Then $f \circ f$ has the power series

$$\begin{aligned} f(f(z)) = 1 &+ c_1^2(z-1) \\ &+ \left[c_1c_2 + c_1^2c_2\right](z-1)^2 \\ &+ \left[c_1c_3 + 2c_1c_2^2 + c_1^3c_3\right](z-1)^3 \\ &+ \cdots . \end{aligned}$$

Comparing coefficients with $1/z = 1 - (z-1) + (z-1)^2 - (z-1)^3 + \cdots$, we find that $c_1^2 = -1$, and $c_2 = 1/(c_1 - 1)$; higher order coefficients, not surprisingly, are underdetermined. At any rate, this shows that

$$f(z) = 1 + (\pm i)(z-1) + \cdots,$$

so, once again, f gives rise to a quarter turn about the point 1.

For the sake of comparison with Section 7, suppose G is a simply connected region, $G = 1/G$, and $1 \in G$. Let $L(z)$ be the branch of the logarithm on G such that $L(1) = 0$, and let $\phi(z)$ be the Riemann mapping function for $L(G)$ that maps 0 to 0; recall that ϕ is an odd function. Put

$$g(z) = \frac{1}{2}\phi^{-1}(L(z))$$

for $z \in G$. Then the function $\sigma(z) = g(z) - g(1/z) = \phi^{-1}(L(z))$ has an inverse, whereupon we may define f by equation (12). It is easily verified that f has property (1) on all of G, not merely on some neighborhood of 1. This brings Theorems 7.2 and 8.1 in line within their common scope.

ACKNOWLEDGMENT. The authors are grateful to P. D. Humke for pointing out [**3**].

REFERENCES

1. R. Anschuetz, H. Sherwood, When is a Function's Inverse Equal to its Reciprocal? *College Math. J.* 27 (1996), 388–393.
2. J. B. Conway, *Functions of One Complex Variable*. Springer-Verlag: New York, 1978.
3. R. Euler and J. Foran, On functions whose inverse is their reciprocal, *Math. Mag.* 54 (1981), 185–189.
4. M. Kuczma, *Functional Equations in a Single Variable*. PWN (Polish Scientific Publishers): Warsaw, 1968.
5. M. Kuczma, B. Choczewski, and R. Ger, *Iterative Functional Equations*. Cambridge University Press: New York, 1990.

RAY CHENG attended the University of Virginia, and received a B.A. in mathematics and physics in 1984. He stayed there until 1989, earning a Ph.D. in mathematics under the guidance of L. Pitt. He then joined the faculty of the University of Louisville. In 1996, just after becoming tenured, and on the eve of his well deserved sabbatical, he left the academic profession to become a corporate executive.
ECI Systems & Engineering, 596 Lynnhaven Parkway, Virginia Beach, VA 23452
rayc@ecihq.com

ABHIJIT DASGUPTA received his Ph.D. in mathematics in 1994 from the University of California, Berkeley, specializing in foundations and real analysis. He taught at the University of Louisville for two semesters in 1994–1995. He joined Advanced Network and Services as a Senior Network Engineer in June, 1995.
2151 Medford Road #34, Ann Arbor, MI 48104
abhijit@ans.net

BRUCE EBANKS received his B.A. from Austin College in 1972 and his Ph.D. from the University of Waterloo in 1977. He was Professor and Mathematics Graduate Director at the University of Louisville when the $1/f$ question arose at the lunch table. Since then he has moved on to become Chairman of the Mathematics Department at Marshall University. In addition to administering and teaching, he enjoys playing the clarinet, running, reading history and science, and doing research involving functional equations. He also observes that West Virginia has great whitewater rafting and hiking.
Marshall University, Huntington, WV 25755
ebanks@MARSHALL.EDU

LAEL KINCH received his Ph.D. from the University of Kentucky in topological semigroups. At the University of Louisville he has co-authored several papers in graph theory and number theory, and counts among his greatest thrills *almost* achieving an Erdős number of 1. He counts this also among his greatest disappointments.
University of Louisville, Louisville, KY 40292
lfkinc01@ulkyvm.louisville.edu

LEE LARSON received his Ph.D. from Michigan State University in 1981, doing real analysis under the watchful eye of C. E. Weil. He was briefly at Western Illinois University before arriving at the University of Louisville in 1983. He was at lunch with Ebanks when the $1/f$ question arose, and shortly after that became Mathematics Graduate Director. In his ever rarer spare time, Larson plays with computers, bicycles, and chess.
University of Louisville, Louisville, KY 40292
llarson@louisville.edu

ROBERT B. McFADDEN was born in Belfast, Northern Ireland. He earned the Ph.D. at the Queen's University of Belfast after entering the university on a classics scholarship. More pear-shaped now, like most former rugby players and coaches, he is a professor of mathematics at the University of Louisville. Before coming to Louisville he worked at QUB, Indiana University, LSU, St. Andrews University in Scotland, Monash University in Melbourne, Australia, Northern Illinois University, and briefly at Argonne National Laboratory. His research interests are in the algebraic theory of semigroups and in automated reasoning. He collects slide rules and makes a wicked martini.
University of Louisville, Louisville, KY 40292
rbmcfa01@ulkyvm.louisville.edu

'Nothing's true,' Postal Weight sobbed, his voice palm-muffled, rocking slightly on the bench.

...

'Todder, you can trust math.'

Freer said 'You heard it here first.'

...

'Todd, trust math. As in Matics, Math E. First-order predicate logic. Never fail you. Quantities and their relation. Rates of change. The vital statistics of God or equivalent. When all else fails. When the boulder's slid all the way back to the bottom. When the headless are blaming. When you do not know your way about. You can fall back and regroup around math. Whose truth is deductive truth. Independent of sense or emotionality. The syllogism. The identity. Modus Tollens. Transitivity. Heaven's theme song. The nightlight on life's dark wall, late at night. Heaven's recipe book. The hydrogen spiral. The methane, ammonia, H_2O. Nucleic acids. A and G, T and C. The creeping inevibatility. Caius is mortal. What it is: listen: it's true.'

'This from a man on academic probation for who knows the length.'

...

'I'm not a math person, Dad says,' said Postal Weight.

...

'The axiom. The lemma. Listen: "If two sets of parametric equations represent the same curve J, but the curve is traced in opposite directions in the two cases, then the two sets of equations produce values for a line integral over J that are negative of each other." Not "*If* thus-and-such." Not "*unless* a gladhanding commercial realtor from Boardman MN in $400 Banfi loafers changes his mind." Always and ever. As in puts the *a* in *a priori*. An honest lamp in the inkiest black, Toddleposter.... Only that at times like this, when you're directionless in a dark wood, trust to the abstract deductive. When driven to your knees, kneel and revere the double S. Leap like a knight of faith into the arms of Peano, Leibniz, Hilbert, L'Hôpital. You will be lifted up. Fourier, Gauss, LaPlace, Rickey. Borne up. Never let fall. Wiener, Reimann, Frege, Green.'

Infinite Jest, by David Foster Wallace;
Boston; Little Brown & Co. (1996) pp. 1071–1072.

Contributed by William Mueller, University of Arizona